1995

HIV, AIDS, Markov Processes, and Health and Disability Insurance

Steven Haberman
City University, London, s.haberman@city.ac.uk

Journal of Actuarial Practice Vol. 3, No. 1, 1995

HIV, AIDS, Markov Processes, and Health and Disability Insurance

Steven Haberman*

Abstract†

This paper presents a Markov model of the transmission and development of HIV and AIDS. The Markov model is used to derive functions needed in the calculation of disability insurance premiums, reserves, and cash flows. An application to health insurance and disability insurance is provided.

Key words and phrases: *permanent health insurance, transition probabilities, premiums, cash flows*

1 Introduction

In the late 1980s the Institute of Actuaries AIDS working party developed a Markov model of the transmission and spread of AIDS among (only) male homosexuals in the United Kingdom; see Daykin et al. (1988a, 1988b, 1990), This model is, in many respects, similar to other mathematical models proposed for the transmission and spread of AIDS.[1]

*Steven Haberman, Ph.D., F.I.A., is professor of actuarial science and dean of the School of Mathematics, Statistics and Actuarial Science at City University, London. He obtained his masters degree in mathematics from Cambridge University and a Ph.D. in actuarial science from City University. Professor Haberman qualified as a Fellow of the Institute of Actuaries in 1975, winning the Burn Prize. He has published widely in actuarial and related fields, and his present research interests include mortality and morbidity modeling, and pension funding.

Dr. Haberman's address is: School of Mathematics, Statistics and Actuarial Science, City University, London ECIV 0HB, England. Internet address: *s.haberman@city.ac.uk*

†This research work was performed under research contract SPES-CT91-0063.

[1]The use of Markov models in actuarial mathematics has been proposed by a number of authors; for example, see Waters (1984, 1989), Amsler (1988), Haberman (1988, 1992), Hoem (1988), Wilkie (1988), Norberg (1988), Ramsay (1989), Continuous Mortality Investigation Report No. 12 (1991), Pitacco (1993), and Jones (1994). For a review of the mathematical models proposed for the transmission and spread of AIDS, see Haberman (1990). Markov models and their applications to life contingencies are included in the Institute and Faculty of Actuaries' syllabus for their professional subject, actuarial mathematics (Subject D). See Chadburn et al. (1993).

The emphasis of actuarial models of HIV/AIDS has been the impact of HIV and AIDS on life insurance underwriting, life insurance premiums, reserving, and, to a lesser extent, health insurance and pension provision. The Institute of Actuaries working party's model attempts (in their own words): "(t)o show the potential impact of HIV on mortality and morbidity and the implications for the use of existing actuarial bases and standard tables for premium rating and reserving". To accomplish this, the Institute of Actuaries working party developed an age-specific model that allows them to use the type of data that normally are available to an insurance company to do the following: (i) to consider the progress of individuals of a given age and gender through future calendar years, (ii) to consider the longer-term trend in transmission, and (iii) to produce numerical results (although not necessarily by analytical means). For these reasons, equilibrium models are of less interest.

The objective of this paper is to develop a modified version of the Institute of Actuaries AIDS working party model. This model then is used to determine the permanent health insurance (PHI) valuation functions needed for the calculation of net premiums and policy values. Examples are given to illustrate formulae and calculations that could be of value in supporting actuarial decisions on pricing and reserving for PHI. In practice, pricing is based on cash-flow models (as well as on present value considerations) using realistic assumptions and allowing for the cost of capital tied up in the establishment of reserves on a more stringent basis. A brief discussion of the equations needed for cash-flow and profit-testing models is provided in Section 6.

The approach advocated in this paper permits the development of partial derivatives of the key valuation functions so that their sensitivities to changes in the underlying parameters (e.g., force of interest, transition intensities) can be measured explicitly. This information is intended to supplement the calculation of sensitivities based on intensive computer-based calculations. An advantage of making these simplified assumptions to the Institute of Actuaries AIDS working party's model is that the resulting simplified model provides an approximation to the transmission of HIV and development of AIDS without the restrictions to the male homosexual population.

The paper is organized as follows: Section 2 provides a brief description of the Institute of Actuaries AIDS working party model. Section 3 describes the modified model used throughout the rest of the paper. Section 4 describes the basics of PHI in the United Kingdom, while Section 5 provides expressions for several PHI valuation functions. Section 6 provides a mathematical description of the expected emerging costs

and cash flows. Finally, Section 7 provides various extensions and modifications to the model described in Section 3.

2 Institute of Actuaries AIDS Working Party Model

The Institute of Actuaries AIDS working party model is a Markov model of the transmission and the progression of HIV among male homosexuals only, with each cohort (of a single age) treated independently. The model assumes that infection occurs from contact between two individuals homosexuals within a single age group.[2] This assumption is artificial, but if infections between those of different ages balance, it may be a reasonable representation of reality. The transition intensities between states are allowed to vary with attained age and time. The model allows for immigration of susceptibles and for normal mortality as well as extra mortality from AIDS.

The AIDS working party made several other simplifying assumptions, including the following:

- That all males described as being at risk of infection behave in the same manner at any one time. As a result, the probability of infection depends on the age of the individual at risk and the particular calendar year, but not on any subdivision according to frequency of sexual contact or frequency of change of sexual partner.

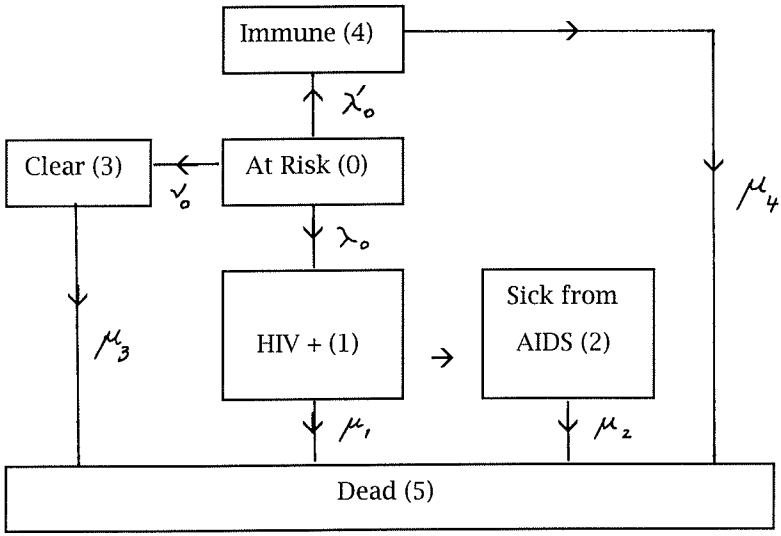
- There are six states. The members of one cohort at age x may be in any one of the six discrete states indicated in Figure 1. Five of these are live states: clear, at risk, immune, positive, and sick from AIDS. The sixth state is the dead state.

  Those in the clear state are those whose sexual activity puts them at no risk of becoming infected with HIV. They form the normal pre-AIDS population for comparative purposes. Those at risk are at risk of acquiring HIV infection through sexual contact with infected persons. Those in the immune state are assumed to have acquired HIV infection and to be infectious, but to be immune from becoming sick from AIDS or dying from AIDS. Those in the positive state are HIV seropositive, but not yet sick from AIDS; they are infectious and not immune.

- It is possible to distinguish those who are HIV seropositive from those who are sick from AIDS. In reality, there are several stages

[2] This model ignores the possibility of heterosexual transmission of the disease and the effects of risk factors such as intravenous drug use and geographic location.

Figure 1
Institute of Actuaries AIDS Working Party Model



in the transition from HIV infection to death from AIDS. Those who are suffering from AIDS are highly infectious, but their sexual activity may be reduced considerably. The model makes it possible to choose whether those sick from AIDS are treated as contributing to further infections or not.

- The current age is part of the status and that transition intensities can vary by current age. In addition, because each age cohort (or year of birth cohort) is treated separately, each transition intensity can be varied by calendar year; therefore, each cohort has its own set of transition intensities. Durations since entry to the states immune, positive, and sick from AIDS are also relevant to the transition intensities.

Possible transitions are as shown in Figure 1. Those in any of the live states may die, and those who are sick from AIDS may die from AIDS or from causes other than AIDS. Those who are at risk may change their behavior and become clear, for example, by giving up sexual activity or by restricting themselves to one equally monogamous partner. There is

no representation in the model of transfer from clear to at risk. Those who are at risk may become infected; at that point, these persons immediately are allocated either to the immune state or to the positive state in proportions that may depend on age (and on calendar year, although it seems unlikely that calendar year would exercise any influence). Those in the positive state may become sick from AIDS, if they do not die first. Infection is possible from the immunes, positives, and sick.

In order not to increase the basic underlying numerical complexity inherent in the model, the Institute of Actuaries working party avoided introducing elements that depend on detailed assumptions about sexual behavior.

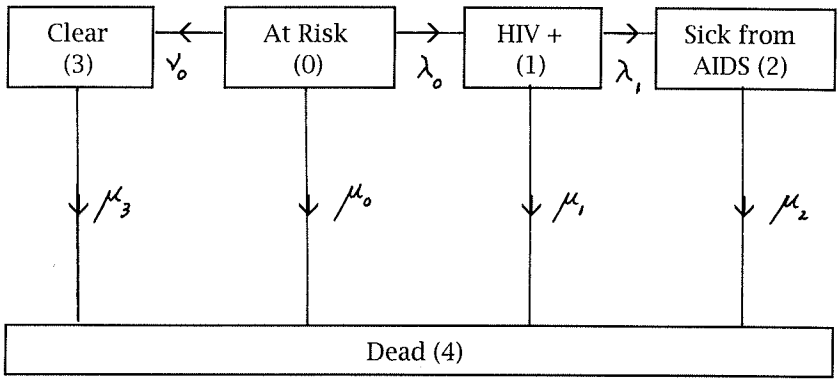
The working party proceeded by establishing a complex series of ordinary and partial differential equations for the probabilities of survival in a state and of transition between states and then solving these equations by numerical methods, given assumptions about the form of the various transition intensities. In most applications, the working party considered the following functional forms for the transition intensities:

- Transition from state 0 to state 1: λ_0 is a function of calendar time, attained age, and number of persons infected;
- Transition from state 0 to state 3: ν_0 is a function of calendar time;
- Transition from state 0 to state 4: λ_0^1 is zero;
- Transition from state 1 to state 2: λ_1 is a function of duration in state 1;
- Transition from state i to state 5: μ_i is a function of attained age for $i = 0$ or 3 or a function of calendar time and attained age for $i = 1, 2$.

3 The Modified AIDS Model

Given the complexity of the Institute of Actuaries AIDS working party model, Haberman (1992) suggested that it be modified along the lines described below so that Markov processes techniques can be applied. Figure 2 depicts the modified version of the model, with the identified states and the corresponding transition intensities. The arrows indicate the directions of the transitions that are permitted in this model.

Figure 2
Modified AIDS Model (Version 1)



The proposed model is a continuous time Markov process with constant transition intensities. Thus, a person in state 0 is subject to a constant force of progression out of state 0 into state 1, out of state 0 into state 3, and to a constant force of mortality out of state 0 into state 4. For a person in state 1, the possible transitions are to states 2 and 4. For a person in state 2 or 3, the transitions are to state 4 only. Once a life leaves a state, it cannot return to that state. The constant transition intensities are as depicted in Figure 2, i.e.,

λ_i = Transition intensity from state i to state $i + 1$, for $i = 0, 1$;
μ_i = Force of mortality in state i, for $i = 0, 1, 2, 3$; and
ν_0 = Transition intensity from state 0 to state 3.

Because the transition intensities remain constant while an individual is in any state, a memoryless property exists. The length of time spent in the current state has no effect on the future length of time that the person will remain in this state.

We now consider how the model of the Institute of Actuaries AIDS working party (as described in Section 2) has been modified to fit with these assumptions.

- First, the immune state is removed.
- Second, the viewpoint is changed from that of the population as a whole to that of an individual male at risk who is considered to progress from state to state over time. We are concerned not with the spread of HIV in a population, but with the outcome for a particular individual.
- Third, it is assumed that all transition intensities are constants, independent of attained age, duration in current state, and secular time. We acknowledge that this assumption contradicts the arguments of Daykin et al. (1988a, 1990) that explain the importance of these variables, in particular attained age, to an actuarial assessment of the effects of HIV and AIDS on survival prospects. Two arguments support this seemingly extreme assumption:
  1. The magnitude of the AIDS-related transition intensities outweighs the normal age-related mortality risk. Many of the AIDS working party simulations assume intensities that do not vary with respect to age;
  2. The desire to reach some analytical results does require, at least initially, some heroic assumptions. We believe that the results are, nevertheless, of value in pricing and reserving.
- Fourth, it is assumed that the transition intensity from the at risk state to the seropositive state is constant and does not depend on the numbers of persons infected. This again is a simplifying assumption to keep the resulting mathematical manipulations tractable.[3] As noted by Daykin et al. (1990), a constant transition intensity from at risk to seropositive would be consistent with the exponential development of new cases of AIDS in the *early* stages of the epidemic.

An advantage to making these simplifying assumptions to the original model is that the model is now flexible enough to approximate the transmission of HIV and the development of AIDS without the restriction to the male homosexual population mentioned in Section 2.

[3]To allow for the effect of heterogeneity of risk and behavioral change, it would be reasonable to postulate an intensity that decreases with time as the epidemic develops. This assumption is not pursued here on the grounds of mathematical tractability.

Given these assumptions, the next step is to determine the transition probabilities. To this end, let $p_{ij}(t)$ be the transition probability that a life now in state i will be in state j at t years from now. There are a number of different ways to set up equations for the required transition probabilities. A common approach used by actuaries (for example, Ramsay (1989), Haberman (1992), and Jones (1994)) is to use the Chapman-Kolmogorov backward system of difference-differential equations. Because the transition intensities are assumed to be constant, we obtain simple recursive solutions to these equations.[4]

Assume that insurance is issued to a life in state i at time of issue, i.e., at $t = 0$. It can easily be proved that

$$p_{00}(t) = e^{-\alpha_0 t} \tag{1}$$

$$p_{11}(t) = e^{-\alpha_1 t} \tag{2}$$

$$p_{22}(t) = e^{-\mu_2 t} \tag{3}$$

$$p_{33}(t) = e^{-\mu_3 t} \tag{4}$$

$$p_{01}(t) = \frac{\lambda_0}{\alpha_0 - \alpha_1}(e^{-\alpha_1 t} - e^{-\alpha_0 t}) \tag{5}$$

$$p_{12}(t) = \frac{\lambda_1}{\alpha_1 - \mu_2}(e^{-\mu_2 t} - e^{-\alpha_1 t}) \tag{6}$$

$$p_{02}(t) = \frac{\lambda_0 \lambda_1}{(\alpha_0 - \alpha_1)(\alpha_0 - \mu_2)(\alpha_1 - \mu_2)}\left[-(\alpha_1 - \mu_2)e^{-\alpha_0 t} + (\alpha_0 - \mu_2)e^{-\alpha_1 t} - (\alpha_0 - \alpha_1)e^{-\mu_2 t}\right] \tag{7}$$

$$p_{03}(t) = \frac{\nu_0}{\alpha_0 - \mu_3}(e^{-\mu_3 t} - e^{-\alpha_0 t}) \tag{8}$$

where $\alpha_0 = \nu_0 + \mu_0 + \lambda_0$ and $\alpha_1 = \lambda_1 + \mu_1$. We note from the representation of the model in Figure 2 that $p_{ij}(t) = 0$ for $i < j$ and that $p_{13}(t) = p_{23}(t) = 0$.

The associated probabilities of dying (being in state 4) are given by the following expression:

$$p_{04} = 1 - \sum_{j=0}^{3} p_{0j}(t)$$

$$p_{14} = 1 - \sum_{j=1}^{3} p_{0j}(t)$$

[4]For a thorough discussion of the Chapman-Kolmogorov backward system of difference-differential equations, see Cox and Miller (Chapter 4, 1965) or Karlin and Taylor (Chapter 4, 1975).

$$p_{24} = 1 - e^{-\mu_2 t}$$
$$p_{34} = 1 - e^{-\mu_3 t}.$$

Alternatively, the equations for $p_{ij}(t)$ can be obtained using matrix methods; see Cox and Miller (Chapter 4.5, 1965).

4 Individual Permanent Health Insurance (PHI)

Individual PHI policies are designed to provide a weekly or monthly income to an individual if he/she is prevented by sickness from working. In this sense, PHI policies provide disability income protection.[5] Policies are typically for a fixed term, usually ceasing at age 65 for males or age 60 for females. Once the insurance company has offered formally to provide the necessary cover and the first premium has been paid, the company cannot cancel the policy as long as the policyholder obeys the policy conditions, hence the name *permanent health insurance.*[6]

Under the most common type of PHI policy, a weekly or monthly income is paid to the policyholder when he/she has been sick for longer than the deferred period.[7] The benefit continues to be paid until the policyholder recovers or dies or until the age at which the policy term ceases. Because the insurer cannot cancel the policy, a policyholder who is sick permanently, or indefinitely, will receive the benefit until one of the above events occurs. With most policies, the premiums are waived while the income benefit is being paid.

Some policies pay a benefit of a fixed level amount, while others provide a benefit that increases to protect the policyholder from inflation. There are various methods by which increases in benefit are provided, some of which are more effective than others.

Another common feature in policy design is a benefit level that reduces with duration of the sickness claim. This is designed to encourage a return to work. The availability of the PHI benefit may lengthen the duration of sickness because of its effect on the minimum acceptable salary that would entice the sick individual to return to gainful employment (i.e., the so-called reservation wage). We allow for the presence of such stepped benefits in our valuation formulae in a later section.

[5] In applying the model to disability insurance policies, it is assumed that at the start of the policy the individual policyholder is in state 0 and healthy and that the insurance company's underwriting or selection process ensures that this is true.

[6] In the U.S. the descriptions *noncancellable* and *guaranteed renewable* are used.

[7] The deferred period is the minimum period of time that the illness must last before the benefit begins. The deferred period is usually four, 13, or 26 weeks.

As PHI policies usually are affected to supplement the sickness benefit available from an employer or the benefit payable from the national or state government (e.g., in the United Kingdom, the national insurance scheme), the deferred period chosen by the policyholder tends to reflect the length of time after which these benefits reduce or (in the case of benefits from an employer) cease. The longer the deferred period, the cheaper the cover and, hence, the lower the premium. Let us now consider the effect of introducing a deferred period of d years on the value of the policy.

For the moment let us assume that a sickness claim is admitted only when full AIDS develops, i.e., the policyholder is in state 2. To calculate present values, we need the probability that an individual starting in state 0 is sick *throughout* the time interval $(t-d, t)$, i.e., that the underlying stochastic process is in state 2 *throughout* this time interval. In the absence of the deferred period, the probability that the policyholder is in state 2 at time t given that the policyholder was in state 0 at time 0 is $p_{02}(t)$. It can be verified by appealing to the Markov property that $p_{02}(t)$ can be written as

$$p_{02}(t) = \int_0^t p_{01}(u)p_{22}(t)\lambda_1 du. \tag{9}$$

We can adapt equation (9) for the presence of a deferred period. Let us define $q_d(t)$ to be the probability that a person in state 0 at time zero is in state 2 throughout the time interval $[t-d, t]$. Adapting the integral definition (9), we can write the following:

$$q_d(t) = \int_0^{t-d} p_{01}(u)p_{22}(t-u)\lambda_1 du, \quad \text{for } t > d. \tag{10}$$

If $t \leq d$, $q_d(t) = 0$. A more formal derivation of $q_d(t)$ is provided by Waters (1984). Such probabilities have been suggested for unemployment insurance by Haberman and Bloomfield (1990) and used extensively by CMIR (1991) for the calculation of PHI-based functions.

To deal with sickness claims that are paid while the policyholder is in state 1 or state 2, let us define $r_d(t)$ to be the probability that a person in state 0 at time zero is in state 1 or state 2 throughout the time interval $[t-d, t]$. Clearly,

$$r_d(t) = \int_0^{t-d} p_{00}(u)(p_{11}(t-u) + p_{12}(t-u))\lambda_0 du, \quad \text{for } t > d. \tag{11}$$

Given the earlier results [equations (3) and (5)], it follows that equation (10) becomes:

$$q_d(t) = \frac{\lambda_0\lambda_1 e^{-\mu_2 t}}{(\alpha_0-\alpha_1)}\left[\frac{1-e^{-(\alpha_1-\mu_2)(t-d)}}{(\alpha_1-\mu_2)} - \frac{1-e^{-(\alpha_0-\mu_2)(t-d)}}{(\alpha_0-\mu_2)}\right]. \quad (12)$$

Similarly, from equations (1), (2), and (6), equation (11) becomes

$$\begin{aligned} r_d(t) &= \frac{\lambda_0}{(\alpha_0-\alpha_1)}\left[\frac{\lambda_1 e^{-\mu_2 t}(1-e^{-(\alpha_1-\mu_2)(t-d)})}{(\alpha_1-\mu_2)}\right. \\ &\quad \left. + \frac{(\mu_1-\mu_2)e^{-\alpha_1 t}(1-e^{-(\alpha_0-\mu_2)(t-d)})}{(\alpha_0-\alpha_1)}\right]. \end{aligned} \quad (13)$$

5 PHI Valuation Functions

Following Daykin et al. (1988b), we recognize that a major difficulty in estimating the impact of HIV infection and AIDS on PHI business is knowing at what stage a PHI claim will be presented to the insurer. For the purposes of illustration of the methodology and the results, we consider here two extreme cases.

Case 1: We assume that a claim only is admitted when full AIDS develops. In the case of a PHI policy with a d week deferred period, we assume that no benefit is payable until d weeks after AIDS has developed (i.e., after entry to state 2).

Case 2: We make the equally extreme assumption that claims are admitted on the basis of HIV seropositivity alone (i.e., on entry to state 1) without requiring evidence of AIDS or any of the intermediate stages.

Let $A_j(n,d)$ be the actuarial present value (under Case j) of a PHI benefit of one unit (per year) in an n year policy (with the n years measured from the inception of the policy) with deferred period of d, for $j = 1, 2$. It follows that:

$$\begin{aligned} A_1(n,d) &= \int_0^n e^{-\delta t} q_d(t)\,dt \qquad \text{Case 1;} \\ &= \lambda_0\lambda_1\left[(a_1+a_2-a_3)e^{-(\delta+\mu_2)d} - a_1 e^{-(\delta+\mu_2)n}\right. \\ &\quad \left. - a_2 e^{-(\alpha_0-\mu_2)d}e^{-(\delta+\alpha_0)n} + a_3 e^{(\alpha_1-\mu_2)d}e^{-(\delta+\alpha_1)n}\right] \quad (14) \\ A_2(n,d) &= \int_0^n e^{-\delta t} r_d(t)\,dt \qquad \text{Case 2;} \end{aligned}$$

$$
\begin{aligned}
= \lambda_0 \Big[& \lambda_1 a_1 (e^{-(\delta+\mu_2)d} - e^{-(\delta+\mu_2)n}) \\
& + a_3(\mu_1 - \mu_2)(e^{-(\delta+\alpha_1)d} - e^{-(\delta+\alpha_1)n}) \\
& - (\lambda_1 a_4 e^{(\alpha_0-\mu_2)d} + a_5(\mu_1 - \mu_2)e^{(\alpha_0-\alpha_1)d}) \\
& \times (e^{-(\delta+\alpha_0)d} - e^{-(\delta+\alpha_0)n}) \Big] \qquad (15)
\end{aligned}
$$

where the a_ks are constants such that their reciprocals are given by

$$
\begin{aligned}
a_1^{-1} &= (\alpha_0 - \mu_2)(\alpha_1 - \mu_2)(\delta + \mu_2) \\
a_2^{-1} &= (\alpha_0 - \mu_2)(\alpha_0 - \alpha_1)(\delta + \alpha_0) \\
a_3^{-1} &= (\alpha_1 - \mu_2)(\alpha_0 - \alpha_1)(\delta + \alpha_1) \\
a_4^{-1} &= (\alpha_1 - \mu_2)(\alpha_0 - \mu_2)(\delta + \alpha_0) \\
a_5^{-1} &= (\alpha_1 - \mu_2)(\alpha_0 - \alpha_1)(\delta + \alpha_0).
\end{aligned}
$$

Given these results, it is possible to investigate the explicit forms for the partial derivatives of $A_1(n,d)$ and $A_2(n,d)$ with respect to d. To illustrate, we present some numerical values for $A_1(n,d)$ and $A_2(n,d)$ based on equations (14) and (15) for different combinations of some of the key parameters.

Following Daykin et al. (1990), we set $\mu_2 = 0.35$ and $\mu_0 = \mu_3 = 0.001$ throughout.[8] We also set $\nu_0 = 0.10$ and $d = 0.07$. Tables 1 to 3 present the magnitudes of $A_1(n,d)$ and $A_2(n,d)$ for the values of λ_0, λ_1, μ_1, n, and d shown. For convenience, we set $\lambda_0 = \lambda_1$ in this presentation. The results indicate that $A_1(n,d)$ and $A_2(n,d)$ both increase with increasing n, decreasing d, decreasing μ_1, and increasing $\lambda_0 = \lambda_1$. They further indicate the relative sensitivities of $A_1(n,d)$ and $A_2(n,d)$ to changes in these parameters and that the ratio $A_1(n,d)/A_2(n,d)$ decreases as $\lambda_0 = \lambda_1$ increases. These results are as expected.

We can compare these results with those given by Daykin et al. (1988b) for the discounted present value of additional sickness benefits under a PHI policy allowing for the two extreme cases described above. Daykin et al. use different morbidity and mortality assumptions (intermediate between the sets underlying Tables 2 and 3). It is impossible to rerun their full model on modified assumptions; however, we can consider from their appendix tables the values of $A_1(n,d)$ and $A_2(n,d)$ and the ratio of the present values under Cases 1 and 2 for comparison with Tables 1 through 3. The details appear in Table 4 for a deferred period of six months and two alternative terminating ages.

[8]The values $\mu_0 = \mu_3 = 0.001$ are approximately equivalent to the value of the force of mortality for a male age 30 to 34 according to English Life Table No 14.

Table 1
Present Values of PHI Benefits: $\lambda_0 = \lambda_1 = 0.001$

Deferred Period: 3 Months Policy Terms (Years)	5	10	15	20
100 A_1	0.00077	0.00341	0.00655	0.00935
100 A_2	.764	2.442	3.673	4.842
Ratio	992	716	561	518
$\mu_1 = 0.01$				
100 A_1	0.00076	0.00332	0.00631	0.00889
100 A_2	0.752	2.172	3.511	4.569
Ratio	989	654	566	514
$\mu_1 = 0.05$				
100 A_1	0.00073	0.00299	0.00538	0.00721
100 A_2	0.700	1.900	2.904	3.599
Ratio	958	635	540	499
Deferred Period: 6 Months Policy Terms (Years)	5	10	15	20
100 A_1	0.00061	0.00293	0.00576	0.00830
100 A_2	0.687	2.127	3.543	4.705
Ratio	1130	726	615	567
$\mu_1 = 0.01$				
100 A_1	0.00061	0.00286	0.00555	0.00790
100 A_2	0.675	2.059	3.381	4.432
Ratio	1110	720	609	561
$\mu_1 = 0.05$				
100 A_1	0.00058	0.00258	0.00474	0.00642
100 A_2	.623	1.788	2.776	3.465
Ratio	1070	693	586	540

The magnitude of $A_1(n,d)$ and $A_2(n,d)$ and the ratios are intermediate between those appearing in Tables 2 and 3 and display similar trends. In particular, we note the stability of the ratios as we consider different age ranges. Similarly, expressions for the present value of premiums and expenses can be developed, including the value of expenses related to the timing of the payment of the sickness benefit. Also, allowance can be made for a waiver of premium benefits and for stepped sickness benefits, i.e., a level of sickness income that depends on the current duration of sickness.

Table 2
Present Value of PHI Benefits: $\lambda_0 = \lambda_1 = 0.01$

Deferred Period: 3 Months Policy Terms (Years)	5	10	15	20
100 A_1	0.0755	0.3257	0.6130	0.8586
100 A_2	7.507	21.567	34.655	44.890
Ratio	99.5	66.2	56.5	52.3
$\mu_1 = 0.01$				
100 A_1	0.0746	0.3178	0.5906	0.8168
100 A_2	7.385	20.911	33.149	42.418
Ratio	99.0	65.8	56.1	51.9
$\mu_1 = 0.05$				
100 A_1	0.0710	0.2863	0.5049	0.6660
100 A_2	6.784	18.320	27.515	33.621
Ratio	96.9	64.0	54.4	50.5
Deferred Period: 6 Months				
Policy Terms (Years)	5	10	15	20
100 A_1	0.0600	.2802	.5396	.7629
100 A_2	6.745	20.458	33.405	43.584
Ratio	112	73.0	61.9	57.1
$\mu_1 = 0.01$				
100 A_1	0.0593	0.2736	0.5201	0.7265
100 A_2	6.625	19.806	31.904	41.117
Ratio	112	72.3	61.3	56.6
$\mu_1 = 0.05$				
100 A_1	0.0566	0.2472	0.4458	0.5937
100 A_2	6.125	17.231	26.288	32.339
Ratio	108	70.0	59.0	54.5

For example, let us take Case 1 (as discussed above, allowing for a waiver of premium benefit) where a sickness claim is admitted only when a transition is made to state 2 (the development of full AIDS). Also, let P_t be the annual premium payable at time $t = 0, 1, \ldots, n-1$. Then the actuarial present value of annual premiums (*APVP*) is given by:

$$APVP = \sum_{t=0}^{n-1} e^{-\delta t} P_t \left(\sum_{j=0}^{3} p_{0j}(t) - q_d(t) \right).$$

Table 3
Present Value of PHI Benefits: $\lambda_0 = \lambda_1 = 0.10$

Deferred Period: 3 Months Policy Terms (Years)	5	10	15	20
100 A_1	6.092	21.527	33.260	40.081
100 A_2	63.091	150.480	206.611	233.975
Ratio	10.4	7.18	6.18	5.84
$\mu_1 = 0.01$				
100 A_1	6.026	20.793	32.223	38.545
100 A_2	62.129	146.437	198.228	224.077
Ratio	10.3	7.04	6.15	5.81
$\mu_1 = 0.05$				
100 A_1	5.745	18.916	28.203	32.804
100 A_2	58.097	130.326	170.037	187.542
Ratio	10.1	6.88	6.03	5.72
Deferred Period: 6 Months Policy Terms (Years)	5	10	15	20
$\mu_1 = 0.001$				
100 A_1	4.895	18.475	29.524	35.867
100 A_2	56.689	142.179	196.820	225.057
Ratio	11.6	7.70	6.67	6.27
$\mu_1 = 0.01$				
100 A_1	4.845	18.081	28.616	34.503
100 A_2	55.748	138.163	189.466	215.189
Ratio	11.5	7.64	6.62	6.24
$\mu_1 = 0.05$				
100 A_1	4.630	16.485	25.088	29.398
100 A_2	51.809	122.173	161.403	178.784
Ratio	11.2	7.41	6.43	6.08

If the deferred period were d_1 and the level of sickness benefit were B per annum for sickness of durations u where $d_1 < u \leq d_1 + d_2$ and the level of sickness benefit were C per annum $(< B)$ for sickness of durations u where $u > d_1 + d_2$, then the actuarial present value of the benefits (*APVB*) under Case i (where $i = 1, 2$) is given by

$$APVB_i = BA_i(n, d_1) + (C - B)A_i(n, d_1 + d_2). \tag{16}$$

Table 4
Present Value of PHI Benefits
According to the Daykin et al. (1988b) Model
Deferred Period Six Months

Terminating Age of Policy: 60 Policy Terms (Years)	10	15	20
Assumption A			
100 A_1	0.67	1.76	2.18
100 A_2	16.77	24.26	26.12
Ratio	25.00	14.00	12.00
Assumptions BC			
100 A_1	0.48	1.02	1.28
100 A_2	9.49	13.88	15.25
Ratio	20.00	14.00	12.00
Assumptions F			
100 A_1	0.65	1.36	1.71
100 A_2	6.88	9.33	10.11
Ratio	11.00	7.00	6.00

Note on assumptions made by Daykin et al. (1988b)

A.	μ_0, μ_1, μ_3	England and Wales population mortality
	λ_0	0.7 at ages 25-50, reducing to zero at ages 15 and 70
	λ_1	Max [exp(-8.4 + 1.4d), 2.5] where d = duration in state 1
	μ_2	Normal mortality +0.7
	ν_0	0
BC.	μ_0, μ_1, μ_3	As for projection A
	λ_0	As for projection A, but reducing linearly from 1987 to 1992 to half initial intensity at all ages
	λ_1	As for projection A
	μ_2	As for projection A
	ν_0	0.10
F.	μ_0, μ_1, μ_3	As for projections A and BC
	λ_0	As for projection BC
	λ_1	As for projections A and BC
	λ_2	Normal mortality +0.35
	ν_0	As for projection BC

Table 4 (cont.)
Present Value of PHI Benefits
According to the Daykin et al. (1988b) Model
Deferred Period Six Months

Terminating Age of Policy: 65 Policy Terms (Years)	10	15	20
Assumption A			
100 A_1	0.34	1.57	2.13
100 A_2	8.81	22.36	25.70
Ratio	26.00	14.00	12.00
Assumptions BC			
100 A_1	0.25	0.90	1.25
100 A_2	4.71	12.13	14.83
Ratio	19.00	13.00	12.00
Assumptions F			
100 A_1	0.38	1.25	1.65
100 A_2	3.96	8.54	9.93
Ratio	10.00	7.00	6.00

Note on assumptions made by Daykin et al. (1988b)

A.	μ_0, μ_1, μ_3	England and Wales population mortality
	λ_0	0.7 at ages 25-50, reducing to zero at ages 15 and 70
	λ_1	Max [exp(-8.4 + 1.4d), 2.5] where d = duration in state 1
	μ_2	Normal mortality +0.7
	ν_0	0
BC.	μ_0, μ_1, μ_3	As for projection A
	λ_0	As for projection A, but reducing linearly from 1987 to 1992 to half initial intensity at all ages
	λ_1	As for projection A
	μ_2	As for projection A
	ν_0	0.10
F.	μ_0, μ_1, μ_3	As for projections A and BC
	λ_0	As for projection BC
	λ_1	As for projections A and BC
	λ_2	Normal mortality +0.35
	ν_0	As for projection BC

6 Emerging Costs

These ideas also can be applied to emerging costs and cash flow considerations. For illustration, we choose a simplified example. We consider a nonprofit PHI policy for a term of n years with a zero deferred period (to simplify the algebra). The policy has annual premiums, P_t, being paid at times $t = 0, 1, \ldots, n-1$. The premium is not paid if the policyholder is sick at time t. We use Case 1 for the definition of sickness for purposes of illustration.

Let us take the benefits provided by the policy to be:

- A death benefit of D_t payable at the end of the policy year if the policyholder dies during the t-th policy year (for $t = 1, 2, \ldots, n$);
- An income benefit if B_t payable at the end of the t-th year if the policyholder is then alive and sick (for $t = 1, 2, \ldots, n$).

We assume that the rate of interest in the t-th policy year is i_t and that the expected cash flow for the t-th policy year per policy alive and in state j at the start of the t-th policy year is $CF_t^{(j)}$. It follows that

$$CF_t^{(j)} = \begin{cases} (1+i_t)P_{t-1} - p_{j2}(1)B_t - p_{j4}(1)D_t & \text{if } j = 0 \text{ or } 1 \\ -p_{22}(1)B_t - p_{24}(1)D_t & \text{if } j = 2 \\ (1+i_t)P_{t-1} - p_{34}(1)D_t & \text{if } j = 3. \end{cases}$$

Then, we define the expected costs (or cash flow) for the t-th policy year per policy originally issued to be EC_t where

$$EC_t = \sum_{j=0}^{3} CF_t^{(j)} p_{0j}(t-1).$$

The above argument can be extended to allow for deferred periods, varying benefits, and transition intensities that are functions of attained age. If the P_ts are net premiums (in the traditional sense of the term), then the equation of value for such a policy would, by definition, be given by

$$\sum_{t=1}^{\infty} e^{-\delta t} EC_t = 0.$$

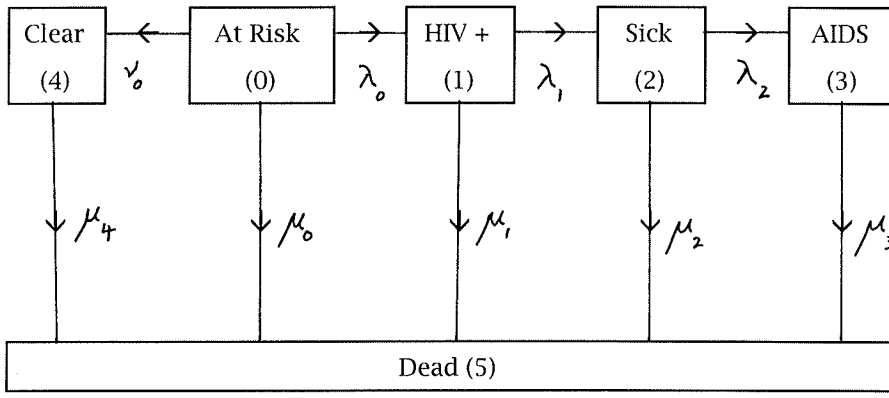
This provides an extension to the results given by Hare and McCutcheon (1991) in respect to conventional life insurance profit testing.

7 Further Modifications to the Model

7.1 Separating Incidence of Disability

The discussion earlier is based on the model depicted in Figure 2 and considers the two extreme cases (Case 1 and Case 2) of timing of a claim for disability income from a PHI policy. A more satisfactory approach is to recognize explicitly the existence of an intermediate state between HIV positive and AIDS for those who are sick and, hence, eligible for a claim. Figure 3 depicts the new model needed, with the states renumbered and the transition intensities as shown. A PHI claim would be accepted once a policyholder has entered state 2 (and the income benefit would be payable while he/she occupies either states 2 or 3).

Figure 3
Modified AIDS Model (Version 2)



This approach leads to no conceptual difficulties. We still must develop the Chapman-Kolmogorov system of differential equations and solve for the transition probabilities $p_{ij}(t)$. The solutions are similar to that given in equations (1) to (6); see Ramsay (1989). For example,

$$p_{22}(t) = e^{-\alpha_2 t} \tag{17}$$

$$p_{23}(t) = \frac{\lambda_2}{\alpha_2 - \mu_3}(e^{-\mu_3 t} - e^{-\alpha_2 t}) \tag{18}$$

where $\alpha_0 = \nu_0 + \mu_0 + \lambda_0$, $\alpha_1 = \lambda_1 + \mu_1$ and $\alpha_2 = \lambda_2 + \mu_2$.

To deal with sickness claims that are paid while the policyholder is in state 2 or state 3 in the presence of a deferred period d, we define $s_d(t)$ to be the probability that a person in state 0 at time 0 is always in state 2 or state 3 during the time interval $[t - d, t]$. Then as for equations (8) and (9), we obtain

$$s_d(t) = \int_0^{t-d} p_{01}(u)(p_{22}(t-u) + p_{23}(t-u))\lambda_1 du, \quad \text{for } t > d. \tag{19}$$

Substitution from equations (5), (17), and (18) leads to the following:

$$\begin{aligned} s_d(t) = {} & \frac{\lambda_0 e^{-\mu_2 t}}{(\alpha_0 - \alpha_1)}\left[\frac{1 - e^{-(\alpha_1-\alpha_2)(t-d)}}{(\alpha_1 - \alpha_2)} - \frac{(1 - e^{-(\alpha_0-\alpha_2)(t-d)})}{(\alpha_0 - \alpha_2)}\right] \\ & + \frac{\lambda_0 e^{-\mu_3 t}}{(\alpha_0 - \alpha_1)(\alpha_2 - \mu_3)}\left[\frac{\lambda_1(1 - e^{-(\alpha_1-\mu_3)(t-d)})}{(\alpha_1 - \mu_3)}\right. \\ & \left. - \frac{\lambda_2(1 - e^{-(\alpha_0-\mu_3)(t-d)})}{(\alpha_0 - \mu_3)}\right] \\ & - \frac{\lambda_0\lambda_2 e^{-\alpha_2 t}}{(\alpha_0 - \alpha_1)(\alpha_2 - \mu_3)}\left[\frac{(1 - e^{-(\alpha_1-\alpha_2)(t-d)})}{(\alpha_1 - \alpha_2)}\right. \\ & \left. - \frac{(1 - e^{-(\alpha_0-\alpha_2)(t-d)})}{(\alpha_0 - \alpha_2)}\right]. \end{aligned} \tag{20}$$

The actuarial present value of a PHI benefit of one unit (per year) in an n year policy with deferred period of d would be:

$$\begin{aligned} A_3(n,d) = {} & \int_d^n e^{-\delta t} s_d(t) dt \\ = {} & \frac{\lambda_0\lambda_1(\mu_2 - \mu_3)(e^{-(\delta+\alpha_2)d} - e^{-(\delta+\alpha_2)n})}{(\alpha_2 - \mu_3)(\alpha_1 - \alpha_2)(\alpha_0 - \alpha_2)(\delta + \alpha_2)} \\ & + \frac{\lambda_0\lambda_1\lambda_2(e^{-(\delta+\mu_3)d} - e^{-(\delta+\mu_3)n})}{(\alpha_2 - \mu_3)(\alpha_1 - \mu_3)(\alpha_0 - \mu_3)(\delta + \mu_3)} \\ & - \frac{\lambda_0\lambda_1(\mu_2 - \mu_3)e^{(\alpha_1-\alpha_2)d}(e^{-(\delta+\alpha_1)d} - e^{-(\delta+\alpha_1)n})}{(\alpha_0 - \alpha_1)(\alpha_1 - \alpha_2)(\alpha_2 - \mu_3)(\delta + \alpha_1)} \\ & + \frac{\lambda_0\lambda_1(\mu_2 - \mu_3)e^{(\alpha_0-\alpha_2)d}(e^{-(\delta+\alpha_0)d} - e^{-(\delta+\alpha_0)n})}{(\alpha_0 - \alpha_1)(\alpha_0 - \alpha_2)(\alpha_2 - \mu_3)(\delta + \alpha_0)} \\ & - \frac{\lambda_0\lambda_1\lambda_2 e^{(\alpha_1-\mu_3)d}(e^{-(\delta+\alpha_1)d} - e^{-(\delta+\alpha_1)n})}{(\alpha_0 - \alpha_1)(\alpha_2 - \mu_3)(\alpha_1 - \mu_3)(\delta + \alpha_1)} \end{aligned}$$

$$+\frac{\lambda_0\lambda_1\lambda_2 e^{(\alpha_0-\mu_3)d}(e^{-(\delta+\alpha_0)d}-e^{-(\delta+\alpha_0)n})}{(\alpha_0-\alpha_1)(\alpha_2-\mu_3)(\alpha_0-\mu_3)(\delta+\alpha_0)}. \tag{21}$$

As an illustration of the numerical effect of separating the incidence of disability and receipt of the income benefit from the onset of AIDS, we present some sample values of $A_3(n,d)$ in Table 5.

Table 5
Present Value of PHI Benefits: Modified Model
Values of 100 A_3

μ_1	μ_2	λ_1	λ_2	5	10	15	20
Deferred Period: Three Months							
Policy Terms (Years)							
0.01	0.05	0.1	0.1	8.702	37.190	66.396	86.776
0.01	0.05	0.1	0.2	8.454	34.488	58.916	74.394
0.01	0.05	0.2	0.1	15.577	60.273	99.365	122.436
0.01	0.05	0.2	0.2	15.121	55.690	87.587	104.136
0.05	0.10	0.1	0.1	7.784	30.408	50.480	62.490
0.05	0.10	0.1	0.2	7.600	28.602	45.928	55.514
0.05	0.10	0.2	0.1	13.962	49.696	76.950	90.890
0.05	0.10	0.2	0.2	13.622	46.619	69.704	80.376
Deferred Period: Six Months							
Policy Terms (Years)							
0.01	0.05	0.1	0.1	7.384	34.163	62.369	82.265
0.01	0.05	0.1	0.2	7.139	31.467	54.898	69.892
0.01	0.05	0.2	0.1	13.289	55.606	93.641	116.331
0.01	0.05	0.2	0.2	12.837	51.032	81.874	98.043
0.05	0.10	0.1	0.1	6.563	27.743	47.066	58.757
0.05	0.10	0.1	0.2	6.380	25.942	42.519	51.788
0.05	0.10	0.2	0.1	11.833	45.521	71.942	85.607
0.05	0.10	0.2	0.2	11.496	42.450	64.705	75.102

We retain the parameter values used in earlier tables ($\mu_0 = \mu_4 = 0.001$, $\nu_0 = 0.10$, $\delta = 0.07$, $\mu_3 = 0.35$); we focus on $\lambda_0 = 0.10$ for direct comparison with Tables 3 and 4. Values have been chosen such that $\alpha_i \neq \alpha_j$ for $i \neq j$. As expected, the values of $A_3(n,d)$ are intermediate between the two extreme estimates of $A_1(n,d)$ and $A_2(n,d)$ presented earlier. We note the extent to which $A_3(n,d)$ increases with increasing n and decreases with increasing d, decreases with increasing forces of mortality, and its relative sensitivity to the choice of λ_1 and relative insensitivity to the choice of λ_2. As expected, $A_3(n,d)$ increases with increasing λ_1 (representing the rate of flow into the claiming state) and decreases with increasing λ_2 (representing part of the rate of flow out of the claiming state). Space constraints prevent pursuing the sensitivities of $A_3(n,d)$ further.

7.2 Dependence on Time of Occupancy

It would be more realistic to allow some transition intensities to depend on the time spent in the current state since the latest transition into that state. This idea of duration dependence leads to the introduction of semi-Markov processes (Cox and Miller, 1965).

The semi-Markov process can be described by a pair of continuous time stochastic processes $\{S(x), Z(x)\}$ for $x \geq 0$. Let $S(x)$ represent the state of an individual at time (or age) x where $S(x) \in 1, 2, \ldots, k$. Let $Z(x)$ denote the duration for an individual at time x of the temporary stay so far in the current state, i.e.,

$$Z(x) = \max\{z : z \leq x \text{ and } S(x-u) = S(x) \text{ for all } u \in [0,z]\}.$$

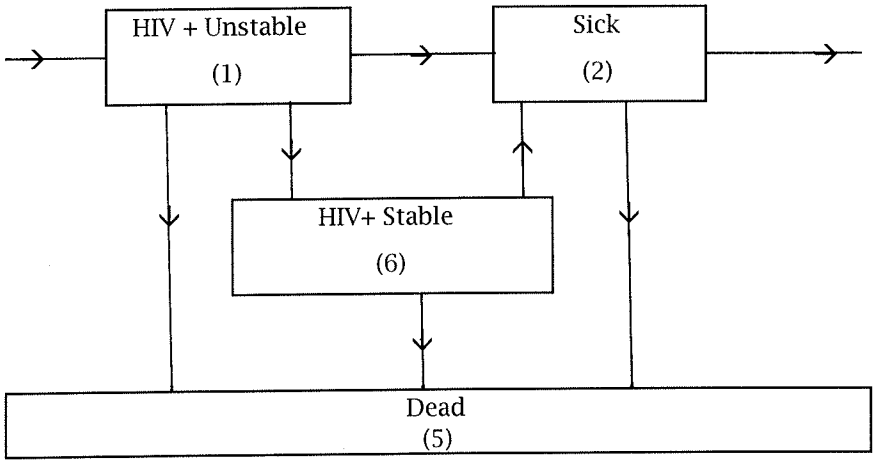
The event $\{S(x) = j \bigcap Z(x) = z\}$ represents an individual being in state j at time x with a duration of z since the last transition into state j.

We can follow the approach of Section 3 and define transition intensities and probabilities and construct equations for the latter (which will be mixed integro-differential equations). The resulting expressions are complex. A useful approximation to the semi-Markov model is to follow the suggestion of Cox and Miller (1965) and introduce a number of substates; this has been applied to actuarial problems by Norberg (1988) in considering select survival models and by Jones (1994) in considering multiple state models.

The replacement of a state by a pair of states labeled "stable" and "unstable" together with the transition intensities that are *independent* of the time spent within each substate mimics approximately the behavior of a semi-Markov model. Figure 4 illustrates part of such a

model structure where we seek to approximate a duration dependence for transition from state 1 to state 2. State 1 becomes "HIV positive: unstable", and we add a further state 6, "HIV positive: stable". This modification can be implemented for the model discussed earlier and depicted in Figures 2 and 3.

Figure 4
Part of Modified AIDS Model (Version 3)



Another approximation that can be made is to model nonconstant intensity functions as piecewise constant functions. This preserves the mathematical tractability of constant forces while giving the flexibility of using nonconstant functions. For a further discussion, readers are referred to Cox and Miller (1965) and to Jones (1994) for actuarial applications.

References

Amsler, M.H. "Sur la Modelisation de Risques Vie Par les Chaines de Markov." *Transactions of 23rd International Congress of Actuaries,* 3 (1988): 1-17.

Chadburn, R.G., Cooper, D.R. and Haberman, S. *Actuarial Mathematics.* Oxford: Institute and Faculty of Actuaries, 1993.

Continuous Mortality Investigation Report No. 12. London: Institute and Faculty of Actuaries, 1991.

Cox, D.R. and Miller, H.D. *The Theory of Stochastic Processes.* London: Chapman and Hall, 1965.

Daykin, C.D., Clark, P.N.S., Eves, M.J., Haberman, S., LeGrys, D.J., Lockyer, J., Michaelson, R.W. and Wilkie, A.D. "The Impact of HIV Infection and AIDS on Insurance in the United Kingdom." *Journal of Institute of Actuaries,* 115 (1988a): 727-837.

Daykin, C.D., Clark, P.N.S., Eves, M.J., Haberman, S., LeGrys, D.J., Lockyer, J., Michaelson, R.W. and Wilkie, A.D. *AIDS Bulletin No. 3.* Report from the Institute of Actuaries AIDS Working Party. London: Institute of Actuaries, 1988b.

Daykin, C.D., Clark, P.N.S., Eves, M.J., Haberman, S., LeGrys, D.J., Lockyer, J., Michaelson, R.W. and Wilkie, A.D. "Projecting the Spread of AIDS in the United Kingdom: A Sensitivity Analysis." *Journal of Institute of Actuaries,* 117 (1990): 95-133.

Haberman, S. "The Central Sickness Rate—A Mathematical Investigation." *Transactions of 23rd International Congress of Actuaries,* 3 (1988): 83-99.

Haberman, S. "Actuarial Review of Models Describing and Predicting the Spread of HIV Infection and AIDS." *Journal of Institute of Actuaries,* 117 (1990): 319-405.

Haberman, S. "HIV, AIDS, and the Approximate Calculation of Life Insurance Functions, Annuities and Net Premiums." *Journal of Institute of Actuaries,* 119 (1992): 345-364.

Haberman, S. and Bloomfield, D.S.F. "Work Time Lost to Sickness, Unemployment, and Stoppages: Measurement and Applications." *Journal of Institute of Actuaries,* 117 (1990): 533-584 (with discussion).

Hare, D.J.P. and McCutcheon, J.J. *An Introduction to Profit Testing.* Actuarial Education Service Special Note. London: Institute of Actuaries, 1991.

Hoem, J.M. "Versatility of the Markov Chain as a Tool in the Mathematics of Life Insurance. Introductory Lecture to Subject 3." *Transactions of 23rd International Congress of Actuaries,* Volume R (1988): 171-202.

Jones, B.L. "Actuarial Calculations using a Markov Model." *Transactions of Society of Actuaries,* 46 (1994): 227-250.

Karlin, S and Taylor, H.M. *A First Course in Stochastic Processes,* 2nd Edition. New York: Academic Press, 1975.

Pitacco, E. "Disability Risk Models: Towards a Unifying Approach." Actuarial Research Paper No 59, City University, London, 1993.

Norberg, R. "Select Mortality: Possible Explanations." *Transactions of 23rd International Congress of Actuaries,* 3 (1988): 215-224.

Ramsay, C.M. "AIDS and the Calculation of Life Insurance Functions." *Transactions of the Society of Actuaries,* 41 (1989): 393-422 (with discussion).

Waters, H.R. "An Approach to the Study of Multiple State Models." *Journal of Institute of Actuaries,* 111 (1984): 363-374.

Waters, H.R. "Some Aspects of the Modeling of Permanent Health Insurance." *Journal of Institute of Actuaries,* 116 (1989): 611-624.

Wilkie, A.D. "Markov Models for Combined Marriage and Mortality Tables." *Transactions of 23rd International Congress of Actuaries,* 3 (1988): 473-486.